Hierarchical and Wavelet-Based Multilinear Models for Multi-Dimensional Visual Data Approximation

Yizhou Yu
Joint work with Qing Wu and Tian Xia
Department of Computer Science
University of Illinois at Urbana-Champaign
Multi-Dimensional Visual Data

- The amount of multi-dimensional visual data is rapidly increasing at an unprecedented rate.

- **Imaging**
  - CCD, images and videos
  - Laser scanning, point clouds
  - Medical Imaging (MRI and DTI), scalar or tensor volume images

- **Simulations**
  - 3D scalar or vector fields in solid and fluid simulations
  - Multi-Dimensional data in appearance modeling and real-time rendering
A Key Challenge

• How can we efficiently
  – Represent
  – Compress
  – Search
  – Analyze
  – Visualize such a vast amount of visual data?

• A possible solution
  – Visual data approximation based on multilinear models
What are Tensors?

• **Multi-Dimensional Matrices**
  - An image ensemble
    \[ \Lambda_{\text{ensemble}} \in \mathbb{R}^{I_{\text{row}} \times I_{\text{col}} \times I_{\text{#images}}} \]
  - A BTF
    \[ \Lambda_{\text{BTF}} \in \mathbb{R}^{I_{\text{row}} \times I_{\text{col}} \times I_{\text{illum}} \times I_{\text{view}}} \]
An Example of BTFs

Illumination

View
Tensor-Matrix Multiplication

Tensor × Matrix Multiplication
Multilinear Approximation

- **Multilinear Models**
  - Rank-$r$ Approximation

\[ \hat{A} = \sum_{j=1}^{r} b_j \times_1 u_j^{(1)} \times_2 u_j^{(2)} \times \cdots \times_N u_j^{(N)} \]

- Rank-$(R_1, R_2, \ldots, R_N)$ Approximation

\[ \tilde{A} = B \times_1 U^{(1)} \times_2 U^{(2)} \times \cdots \times_N U^{(N)} \]
Multilinear Approximation

- Rank-\((R_1, R_2, \ldots, R_N)\) Approximation

\[ A \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}, \quad \tilde{A} = \arg \min_{\hat{A}} \| A - \hat{A} \|^2 \]

\[ \tilde{A} = B \times_1 U^{(1)} \times_2 U^{(2)} \times \cdots \times_N U^{(N)}, \quad U^{(i)} \in \mathbb{R}_i \times R_i, \quad i \in (1, N) \]

A special case:
\[ C \in \mathbb{R}^{I_1 \times I_2}, \quad SVD(C) = USV^T = S \times_1 U \times_2 V \]

- Alternative Least Squares (ALS)

- Initialization
  - N-mode SVD vs. \[ U^{(n)}_0 = \begin{bmatrix} I_{R_n} & 0 \end{bmatrix}^T \] (random numbers)
Related Methods

• **PCA and tensor approximation** (*adaptive bases*)
  - [Kroonenberg and Leeuw 1980]
  - [Lathauwer et al 2000], [Wang et al 2005]

• **Wavelet analysis** (*fixed bases*)
  - Wavelet analysis [Antonini et al 1992], [Muraki 1993]
  - Wavelet packet [Meyer et al 2000]
  - Oriented wavelet bases
Applications

- Precomputed Radiance Transfer and Real-Time Rendering

[Tsai and Shih 2006]
Applications

• Nearest Neighbor Search
Tensor Construction

• **Datum (Image) as Is**

  – An image ensemble
  \[ \Lambda_{ensemble} \in \mathbb{R}^{I_{row} \times I_{col} \times I_{#images}} \]

  – A BTF
  \[ \Lambda_{BTF} \in \mathbb{R}^{I_{row} \times I_{col} \times I_{illumin} \times I_{view}} \]

• **Datum (Image) as Vector**

  – Traditional PCA

  – TensorTexture, SIGGRAPH 2004

  \[ M_{ensemble} \in \mathbb{R}^{(I_{row}I_{col}) \times I_{#images}} \]

  \[ A_{BTF} \in \mathbb{R}^{(I_{row}I_{col}) \times I_{illumin} \times I_{view}} \]

[ Wang et al., SIGGRAPH 2005 ]
Advantages of Datum-as-Is

- Spatial locality and redundancy

It encodes:
1. Pixel-wise covariance
2. Row/column covariance
Rank-1 Basis Images
Level-2 Basis Images for the LEGO BTF

Reconstructed Images
RMSE Errors: Lichen and Velvet

![Graph showing RMSE errors for Lichen and Velvet with different compression ratios.]
Original | PCA | Modified TensorTexture
---|---|---
TensorTexture

Our Method
BTF Mapping and Rendering

Vase and Sponge

Textures on the vase and the sponge are reconstructed from compressed BTF tensors
(1 strong moving point light source and 2 weak static ones)
Characteristics of Visual Data

- Multiple scales
- Spatially inhomogeneous

Fourier Transform

- Second harmonic
- First harmonic
- Fundamental
- Composite
Hierarchical Transformation and Approximation

- **Top-down hierarchical transformation (lossless)**
  - Tensor (ensemble) approximation at the current level
  - Subdivide every residual tensor into $2^N$ smaller tensors by bisecting each dimension
  - Repeat until the subdivided tensors are sufficiently small

[Wu et al. TVCG 2008]
Hierarchical Transformation and Approximation

• **Lossy approximation**
  - Pruning residual tensors when their energy is below a threshold
  - Quantization on core tensor coefficients
    • 8-20 bits per coefficient

• **How to determine the reduced ranks at each level?**
  - Optimal solution is expensive
  - Choose \( r_1, r_2, \ldots, r_N \) for the top level
  - On subsequent levels, each of the ranks follows a geometric progression.
Progressive Approximation

different levels of approximation
Correlation among Residual Tensors

- Strong correlation among subdivided tensors
- Correlation among color channels and vector components
Tensor Ensemble Approximation

• **Collective tensor approximation**
  - Rank-$(R_1, R_2, \ldots, R_N)$ approximation of $A_1, A_2, \ldots, A_m \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$
  - Pile $A_1, A_2, \ldots, A_m$ into a $(N+1)$-th order tensor $G$
  - Perform rank- $(R_1, R_2, \ldots, R_N)$ approximation of $G$ ($r < m$)
Tensor Ensemble Approximation

• (Sub-)tensors share basis matrices to reduce basis overhead and improve compression efficiency

original

individual
PSNR 20.13

ensemble
PSNR 26.17

87.5% compression
Experiments and Results

• **Input Data**
  – 4D BTFs
    • *Sponge, lichen*
  – 3D medical images in the Visible Human dataset
    • *Head, abdomen*
  – 4D time-varying scientific dataset
    • *Velocity and energy field*

• **Compression ratio and PSNR**
  – Bases are overheads
  – Small wavelet coefficients are discarded
Experimental Results

<table>
<thead>
<tr>
<th>Original</th>
<th>Single-level</th>
<th>Hierarchical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR 29.89</td>
<td>PSNR 31.05</td>
</tr>
</tbody>
</table>

93.6% compression
Experimental Results

(a) Original

(b) Wavelet
98.2% Compression
PSNR 16.77

(c) Single-Level
98.2% Compression
PSNR 24.05

(d) Multi-Level
98.2% Compression
PSNR 25.21

(e) Wavelet Packet
98.2% Compression
PSNR 20.54

(f) Single-Level
99.975% Compression
PSNR 20.11

(g) Multi-Level
99.975% Compression
PSNR 21.01
Experimental Results

(a) Original
(b) Wavelet
PSNR 35.12
(c) Single-level
PSNR 43.56
(d) Multilevel
PSNR 45.41
(e) Wavelet Packet
PSNR 39.48
(f) Residual of (c)
(g) Residual of (d)
Experimental Results

Visible Human

Rocket Jets Simulation
Conclusion and Limitation

• For high-dimensional datasets, the gained efficiency surpasses the basis overhead

• For 2D images, the basis overhead cannot be tolerated

Multilinear models suffer basis overhead and sharp-edge smoothing for 2-D images. The resolution of LENA is 256x256 and the three color channels are processed as an ensemble. A rank-(32,32) approximation imposes $\frac{2(256^2)}{3(32^2)+2(256^2)} = 84.21\%$ overhead at 90.1% compression. Furthermore, it blurs sharp edges compared with wavelet transform.
Wavelet-Based Hybrid Multilinear Models

- **Multilinear Models**
  - Compact approximation with optimal adaptive bases
  - Large basis overhead for low-dimensional images

- **Wavelet (Packet) Transforms**
  - Suboptimal fixed bases
  - No basis overhead
  - Feature sensitive filters
  - Sparse coefficients

- **Hybrid models**
  - Clustered tensor approximation of sparse wavelet coefficients
  - [Wu et al., ICIP 2008]
Overview

• Hybrid Multilinear Models in the Wavelet Domain
  – Subdivide high-frequency subbands so that most blocks get pruned
  – Group remaining blocks into clusters
  – Approximate each cluster as a tensor ensemble
Subband Data Organization

- **Block Subdivision**
  - High-frequency subbands are usually very sparse
  - By subdividing them into small blocks we can discard most blocks and focus on those with a significant energy
  - We use 2x2 blocks in our experiments

- **Block Flattening**
**Subband Data Organization**

- **Channel Stacking**
  - Different color channels have strong correlation
- **Subband Stacking**
  - Same-level high-frequency subbands have strong correlation
- **Organized Tensors**
  - A collection of tensors for each level
  - Always 3rd order: spatial, subbands, color channels (could be degenerate)
Subband Data Organization

- Wavelet Packet Transform
Tensor Clustering

• An EM algorithm
  – Initialization
    • Random
    • Use GPCA result
  – Iteratively refine clusters
    • Update basis matrices for every cluster
    • Project each tensor into every cluster using its basis matrices and update the cluster membership of the tensor according to projection errors
  – Stop when clusters do not change
Rank Optimization

- **Tensor Approximation**
  - How to determine the ranks?
    \[
    A_{n_b \times n_s \times n_c} = B_{n_b \times n_s \times n_c} \times_1 U^{(b)}_{n_b \times n_b} \times_2 U^{(s)}_{n_s \times n_s} \times_3 U^{(c)}_{n_c \times n_c}
    \]
    \[
    MSE = \frac{1}{n_b n_s n_c} \left( \sum_{i=1}^{n_b} \sum_{j=1}^{n_s} \sum_{k=1}^{n_c} B_{ijk}^2 - \sum_{i=1}^{r_b} \sum_{j=1}^{r_s} \sum_{k=1}^{r_c} B_{ijk}^2 \right) \leq \epsilon^2
    \]
  - Search for the most economic ranks while controlling MSE
    \[
    (r_b, r_s, r_c) = \text{argmin } n_b r_b + n_s r_s + n_c r_c + r_b r_s r_c, \text{s.t.}
    \]
    \[
    \sum_{i=1}^{r_b} \sum_{j=1}^{r_s} \sum_{k=1}^{r_c} B_{ijk}^2 \geq \sum_{i=1}^{n_b} \sum_{j=1}^{n_s} \sum_{k=1}^{n_c} B_{ijk}^2 - n_b n_s n_c \epsilon^2,
    \]
    \[
    1 \leq r_b \leq n_b, 1 \leq r_s \leq n_s, 1 \leq r_c \leq n_c.
    \]
Rank Optimization

- Tensor Ensemble Approximation

\[ A \approx B \times_1 U^{(b)}_{n_b \times r_b} \times_2 U^{(s)}_{n_s \times r_s} \times_3 U^{(c)}_{n_c \times r_c} \times_4 U^{(t)}_{n_t \times r_t} \]

\[
(r_b, r_s, r_c, r_t) = \text{argmin} \quad n_b r_b + n_s r_s + n_c r_c + n_t r_t + r_b r_s r_c r_t, \text{ s.t.} \]

\[
\sum_{i=1}^{r_b} \sum_{j=1}^{r_s} \sum_{k=1}^{r_c} \sum_{h=1}^{r_t} B_{ijkh}^2 \geq \sum_{i=1}^{n_b} \sum_{j=1}^{n_s} \sum_{k=1}^{n_c} \sum_{h=1}^{n_t} B_{ijkh}^2 - n_b n_s n_c n_t \epsilon^2
\]

\[ 1 \leq r_b \leq n_b, 1 \leq r_s \leq n_s, 1 \leq r_c \leq n_c, 1 \leq r_t \leq n_t. \]
Experiments and Results

• 2D Images
  – SCENE, LENA, BABOO

• 3D Visible Human Medical Datasets
  – From US National Library of Medicine
  – HEAD, ABDOMEN
Experimental Results

Input

Multilinear, PSNR=19.18
97.5% compression

Wavelet, PSNR=20.95

Wavelet Packet, PSNR=22.12

Wavelet+HM, PSNR=22.01

Wavelet Packet+HM, PSNR=23.11
Experimental Results

Input

Wavelet, PSNR=26.95

Wavelet Packet, PSNR=27.95

Wavelet Packet+HM, PSNR=30.21

Wavelet+HM, PSNR=30.18

98% compression
Experimental Results

Input

Wavelet, PSNR=24.95

Wavelet Packet, PSNR=26.18

Wavelet+HM, PSNR=27.23

Wavelet Packet+HM, PSNR=28.05

98% compression
Experimental Results

Input

Multilinear PSNR=26.44

Wavelet Packet PSNR=29.69

Wavelet Packet+HM PSNR=30.75

99.5% compression
Experimental Results

- baboo
  - Wavelet
  - Wavelet Packet
  - Wavelet + Tensor
  - Wavelet Packet + Tensor

- abdomen
  - Wavelet
  - Wavelet Packet
  - Wavelet + Tensor
  - Wavelet Packet + Tensor
Acknowledgments

- UIUC: Qing Wu, Tian Xia, Hongcheng Wang, Narendra Ahuja, Lin Shi
- NSF
- UIUC Research Board
Thank You!